Data Analysis of Data Challenge 4:
The SRD and Science Commissioning

James Annis
Center for Particle Astrophysics
Fermi National Accelerator Laboratory

Data Quality is Paramount

1. Data quality is paramount
   - Survey requirements are defined by the core science goals.
   - Design the instrument and survey strategy that will meet these goals.
   - Determine what the instrument, the strategy, and the laws of physics allows
   - Write the data reduction software so that it extracts the highest quality it can
   - Design quality validation tools to verify that the data meet these requirements.

2. Data quality is paramount despite knowing that large projects are always tight on time and money and the temptation is to skimp on data quality.
   - It is cheaper to do things right the first time.

3. This is the first in what should be a yearly report on progress of the Data Challenges versus the SRD.

4. Talk Outline
   - Science Requirements Documents
   - Evaluation of DC4 data against the requirements
   - Evolution of these tests into the tests of science commissioning

Science Requirements Documents

1. The survey requirement documents

   - DES Science Requirement Document (SRD)
   - DESDM Requirements and Technical Specifications document
   - DECam Technical Specifications and Requirements
   - DECam Community Pipeline Software Requirements and Technical Specifications

2. The Science Requirement Document

   - Installed as a survey document in 2004.
     - That version aimed at defining requirements of an imaging survey useful for cosmology.
   - The formation of the Science Committee and working groups brought new intellectual strength to the collaboration. The London Process sought to harness this strength by requesting requirement documents from each working group. The Ad-Hoc Committee on Science Requirements, chaired by J. Frieman, was formed in Sept 2007 to integrate these documents into a rewritten SRD.
     - This version aimed to set the requirements in terms of effect on the DETF figure or merit.
   - The version of the SRD on the review page (v9.1) is the first document from the Committee. Newer drafts exist (e.g. v9.4a), incorporating commentary and new calculations, but it is not in a form ready to be reviewed.
     - v9.1 was reviewed by the Joint DOE/NSF Review of DES in Jan 2008 and termed a mature document. It was recommended that we finalize the document quickly.

3. The London Process must be brought to a close, and a new SRD signed off, so that a new round of flow downs to the other documents can begin.

Testing DC4 Against Science Requirements

1. Given that there exists requirement documents, it is of interest to know how we are doing in meeting the requirements.

2. We report on testing the DC4 outputs against requirements.
   - This is good news: the DC4 data is the result of a system mature enough to test.

3. We will draw requirements from both the Science Requirements Document and from the DESDM Requirements and Technical Specifications.

4. We report on a subset of the requirements
   - Image quality
   - Limiting magnitude
   - Astrometric calibration
   - Photometric calibration
     - Galaxy catalog content
   - Photometric redshifts

5. A goal is to develop tests that will be useful as science commissioning tests. Hence a preference is given to data driven tests rather than truth tables.

Image Quality

1. The median PSF FWHM (averaged over all exposures and over the survey area) in each of the *r, i,* and *z* bands should be less than 0.9" ✕
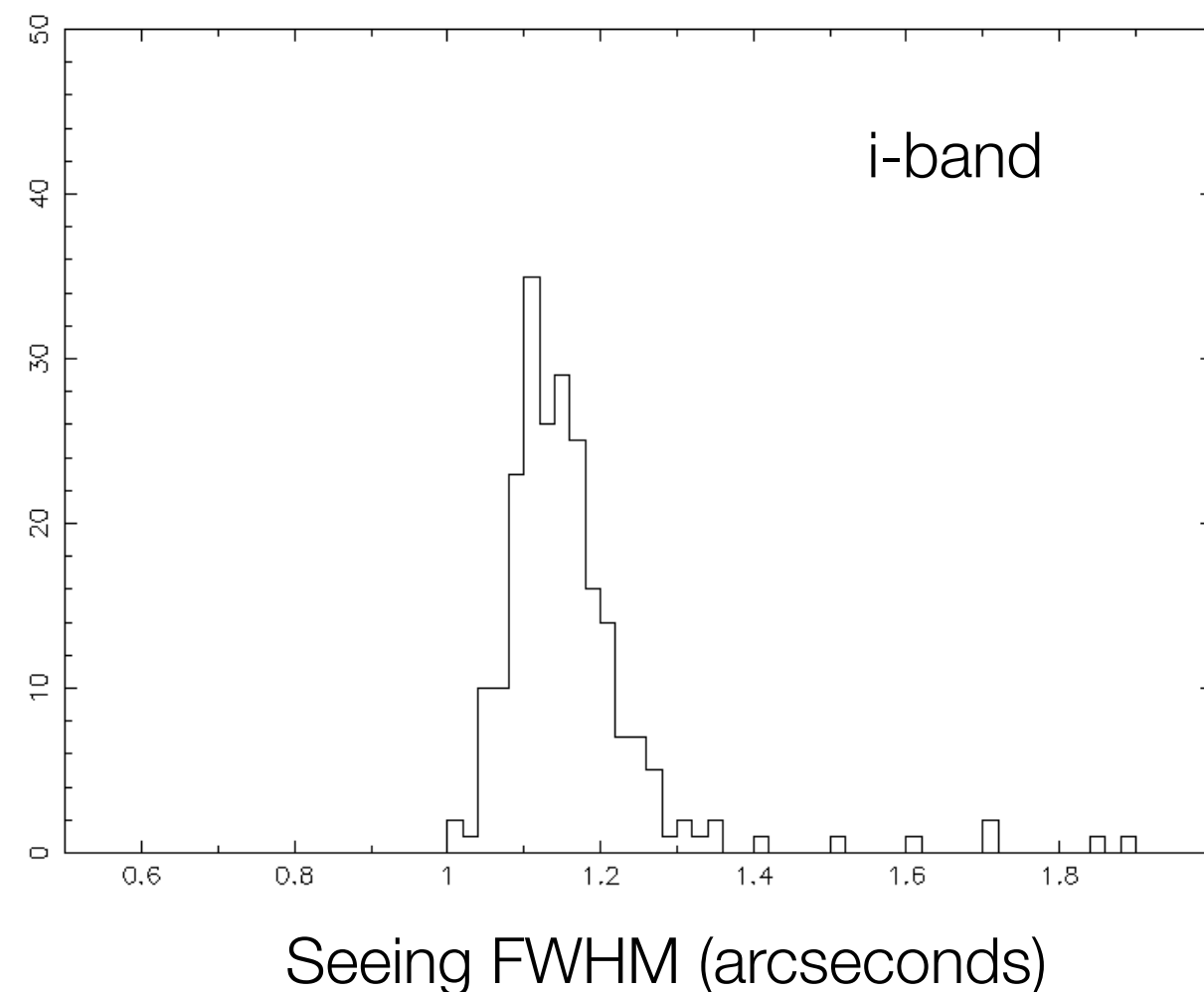
2. fwhm_world from $18 < i < 20$ stars

3. Median FWHM
    - g: 1.3"
    - r: 1.2"
    - i: 1.1"
    - z: 1.1"
    - Y: 1.2"

4. This is likely a simulation issue.

5. Side effect is to make seeing issues worse in DC4

i-band

Seeing FWHM (arcseconds)

Limiting Magnitude I

1. The requirements are grizY= 24.0, 24.0, 24.0, 23.6, 21.6 ✓✗ (see next page)

   - The catalog should provides useful data to these limits

2. Cumulative exposures time

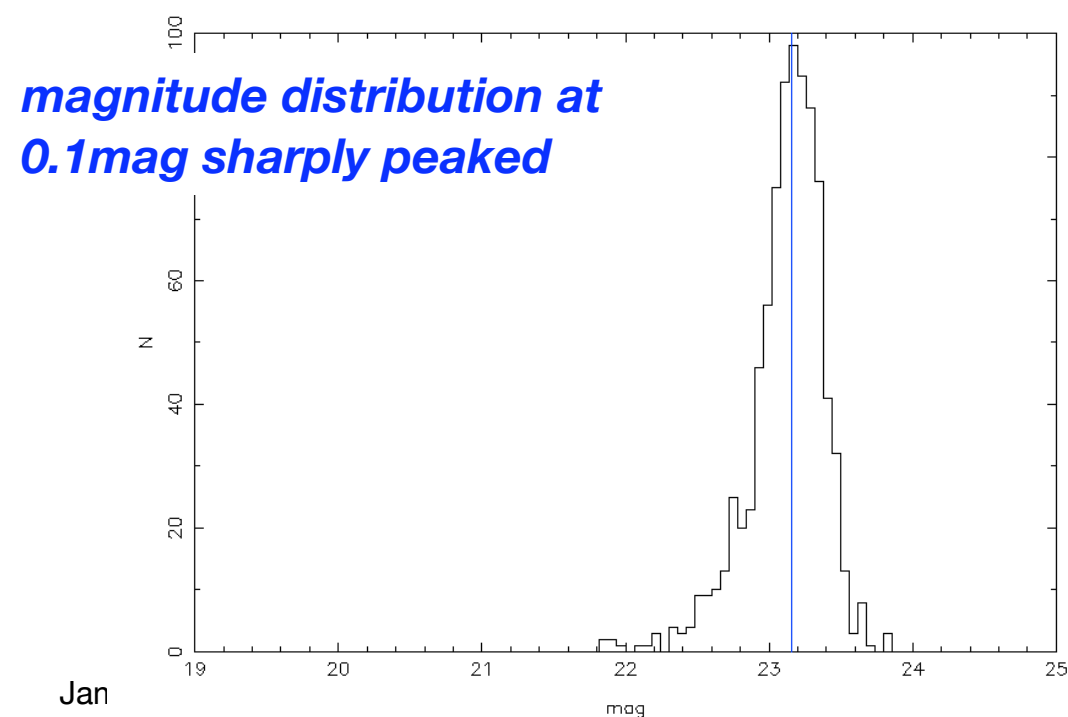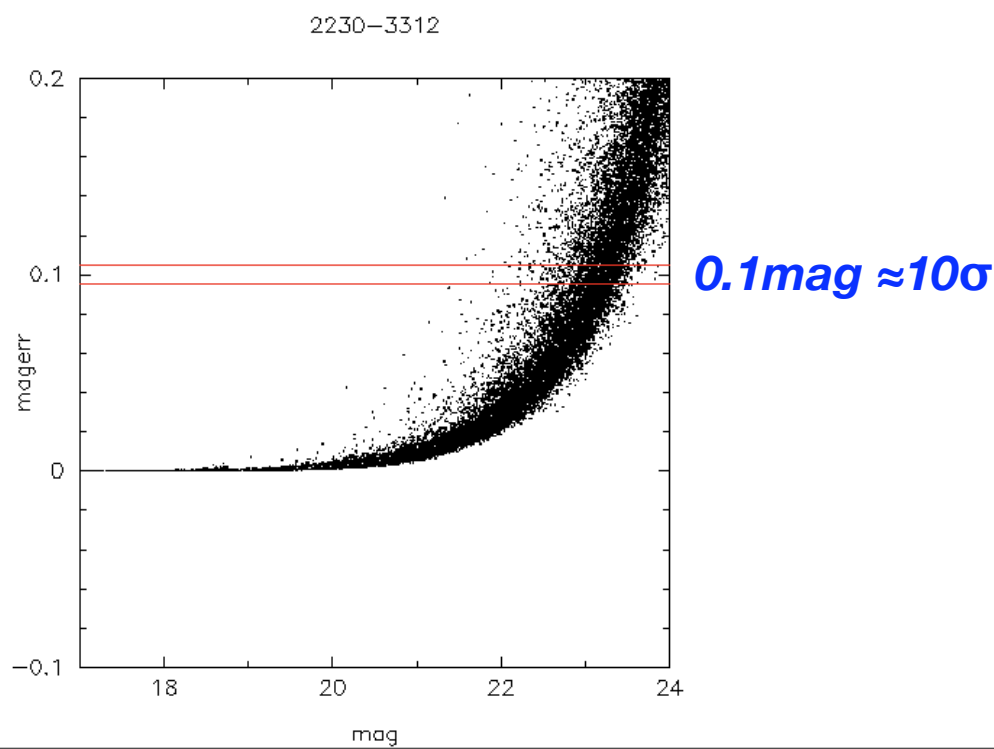   - Current survey strategy: time(grizY) = 320 320 1120 1920 320
   - DC4 simulation: time(grizY) = 800 800 1500 1500 600

   DC4 depths differ from plans- recalculate depths

3. Predicted depth in data:

   - mag (grizY; 10σ in 1.5" aperture) = 25.1 24.7 24.6 23.8 21.8

4. Measure limiting magnitude by locating 0.1 mag error point in catalog

0.1mag ≈10σ

magnitude distribution at 0.1mag sharply peaked

Limiting Magnitude II

1. Types of magnitudes:
   - aperture magnitudes ✓   - Total magnitudes (e.g. SDSS Petrosian mags) ✓
   - PSF magnitudes ✗   - Color optimized magnitudes (e.g. SDSS cModel mags) ✗

2. Catalog has aperture magnitudes and mag_auto
   - Assume mag_aper1 is used as color magnitude, mag_auto as total magnitude

3. mag_aper1 (1.5" aperture)
   - grizY (measured) = 25.2 24.9 24.7 23.8 22.0
   - mag-mag$_{predicted}$ = -0.1 -0.4 -0.2 -0.1 -0.3
   - The catalog is as deep as expected.

4. mag_auto (elliptical aperture, diameter set by isophotal 2nd moment*factor)
   - predicted is 1.5" aperture, corrected for 0.3 mags galaxy light outside aperture
   - grizY (measured) = 23.3 22.9 22.8 21.9 19.9
   - mag-mag$_{predicted}$ = 1.5 1.5 1.5 1.6 1.6
   - Ideally this would be closer to 0. The big aperture radius of mag_auto causes a noisy measurement. Very akin to the behavior of the SDSS Petrosian magnitude.
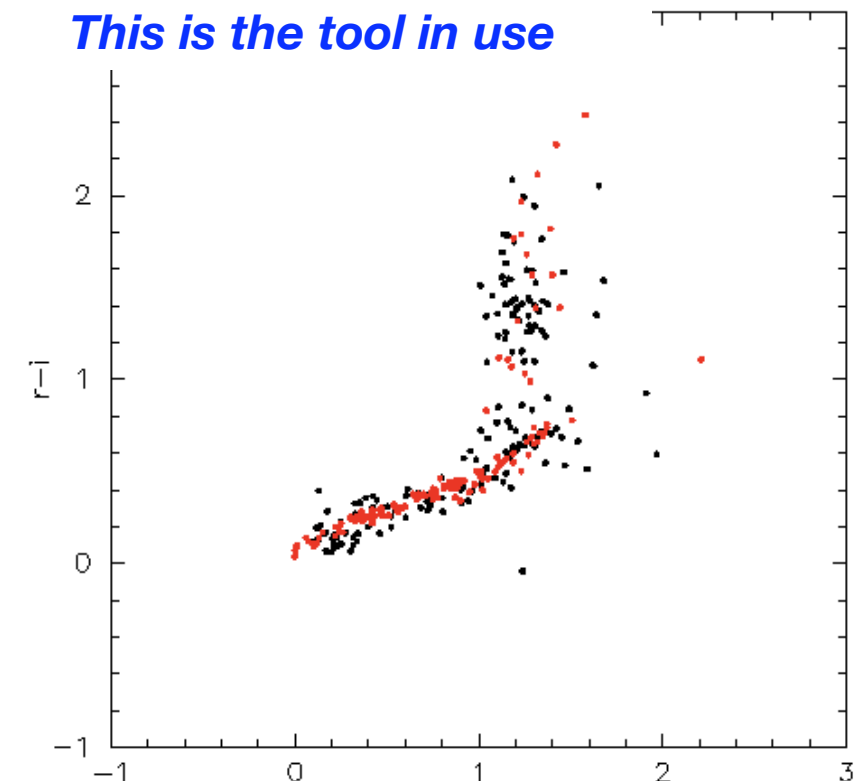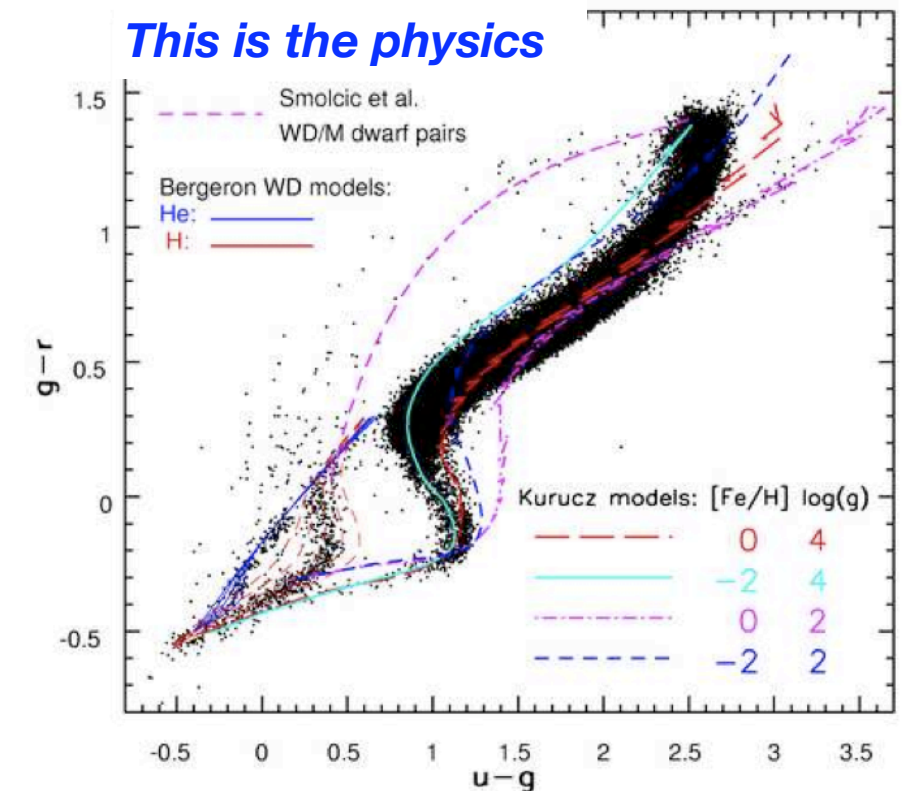
Astrometric Calibration

1. Astrometry no worse than 100 mas✓

2. Test using USNO-B stars and truth tables

3. Plots show residuals over survey area
   - means and unclipped RMS
   - 0.1 degree bins
   - 1000-2000 stars in each
   - 267,000 stars total

4. Mean, an astrometric frame shift, near 0

5. Dispersion, noise in astrometric solution, ~60 mas.

6. This is good performance.

dc4 rband – USNO truth comparison
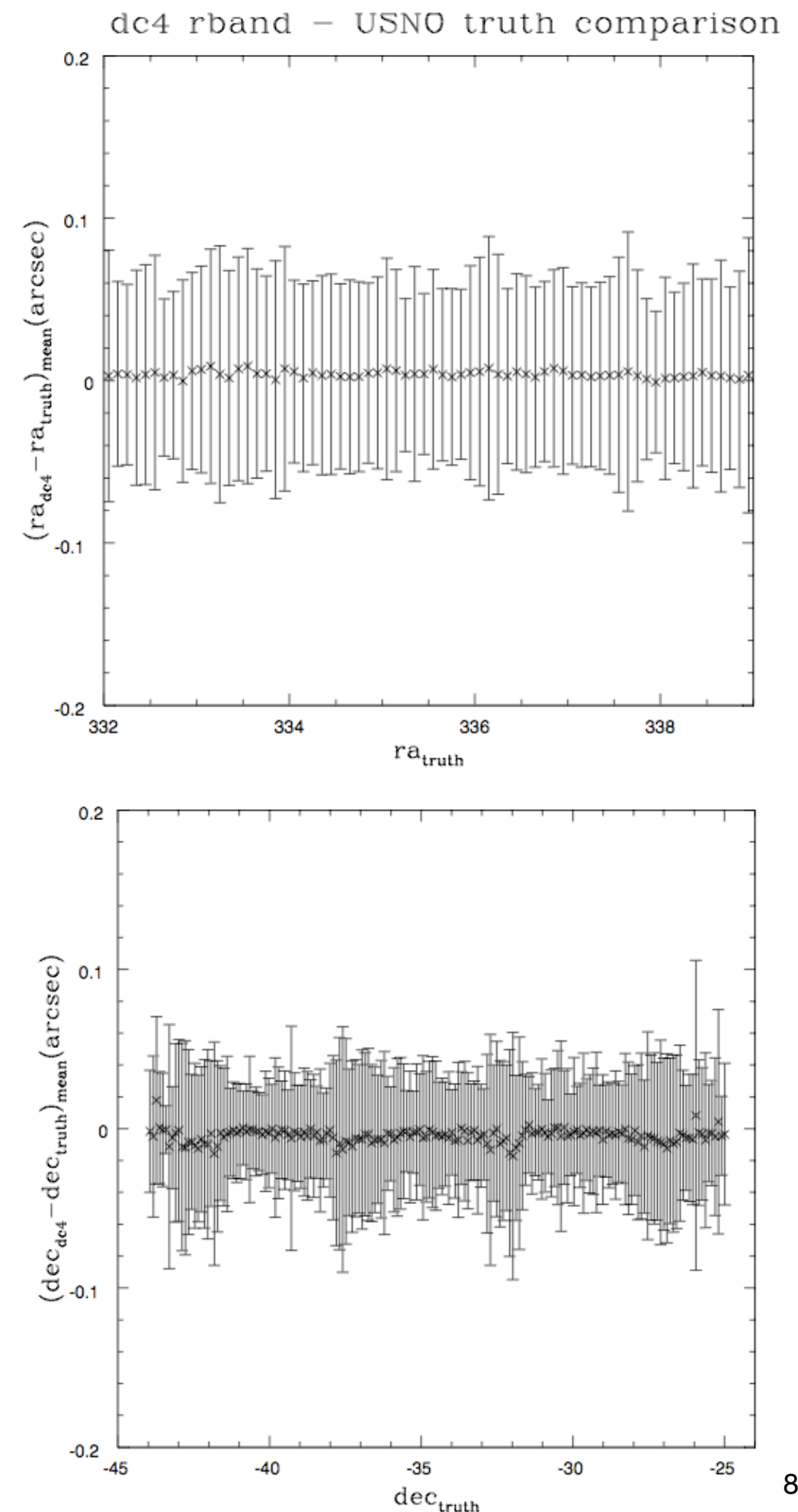
Photometric Calibration I

1. We will test this using the colors of stars in the simulation rather than truth tables.

2. The stellar locus in color-color space has very small intrinsic scatter.
   - e.g., Ivezic et al 2007, who uses it to test calibration

3. We will fit the stars in DC4 with a stellar locus
   - stellar locus computed using a subset of the Gunn-Stryker stellar spectral atlas, transformed to DES colors.

4. Star-galaxy separation must use the position dependent model of the PSF and be accurate at >= 95% at the 10 sigma photometric limit.✕
   - STAR_CLASS becomes uncertain at i > 21.

Stripe 82 star magnitudes

This is the physics

This is the tool in use

Red: Gunn-Stryker stars, DES colors
Black: Stars in tile 2234-3145 at 20<i<22.

James Annis *Directors' Review*
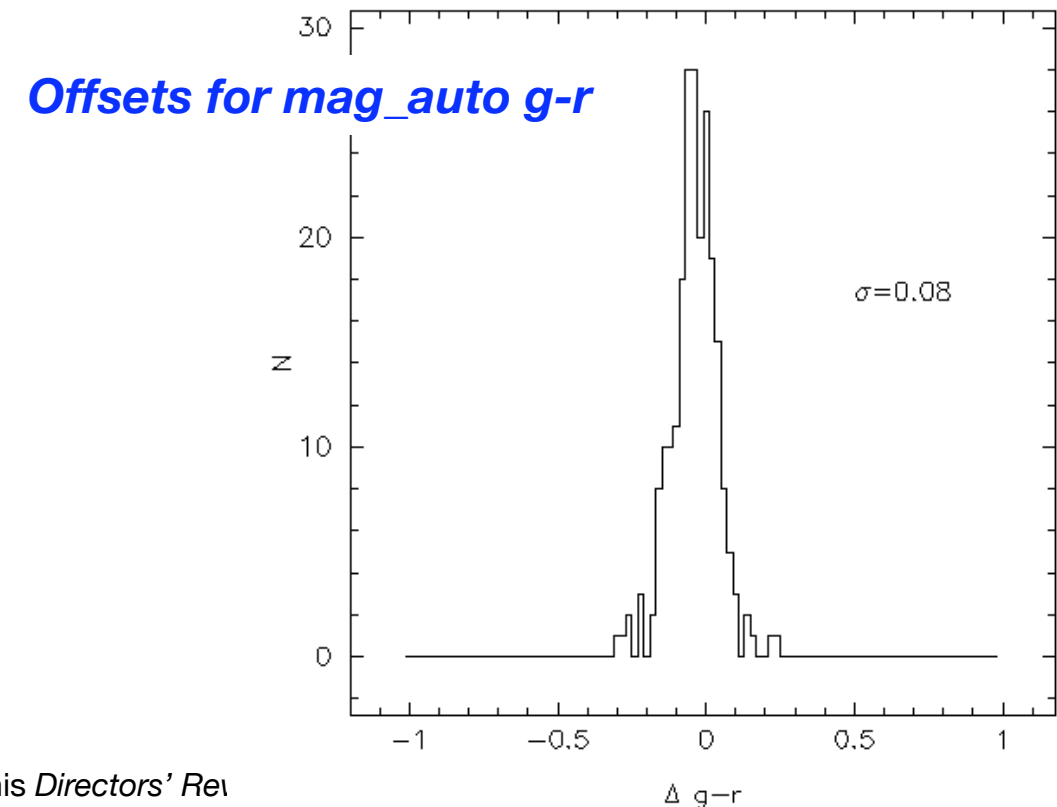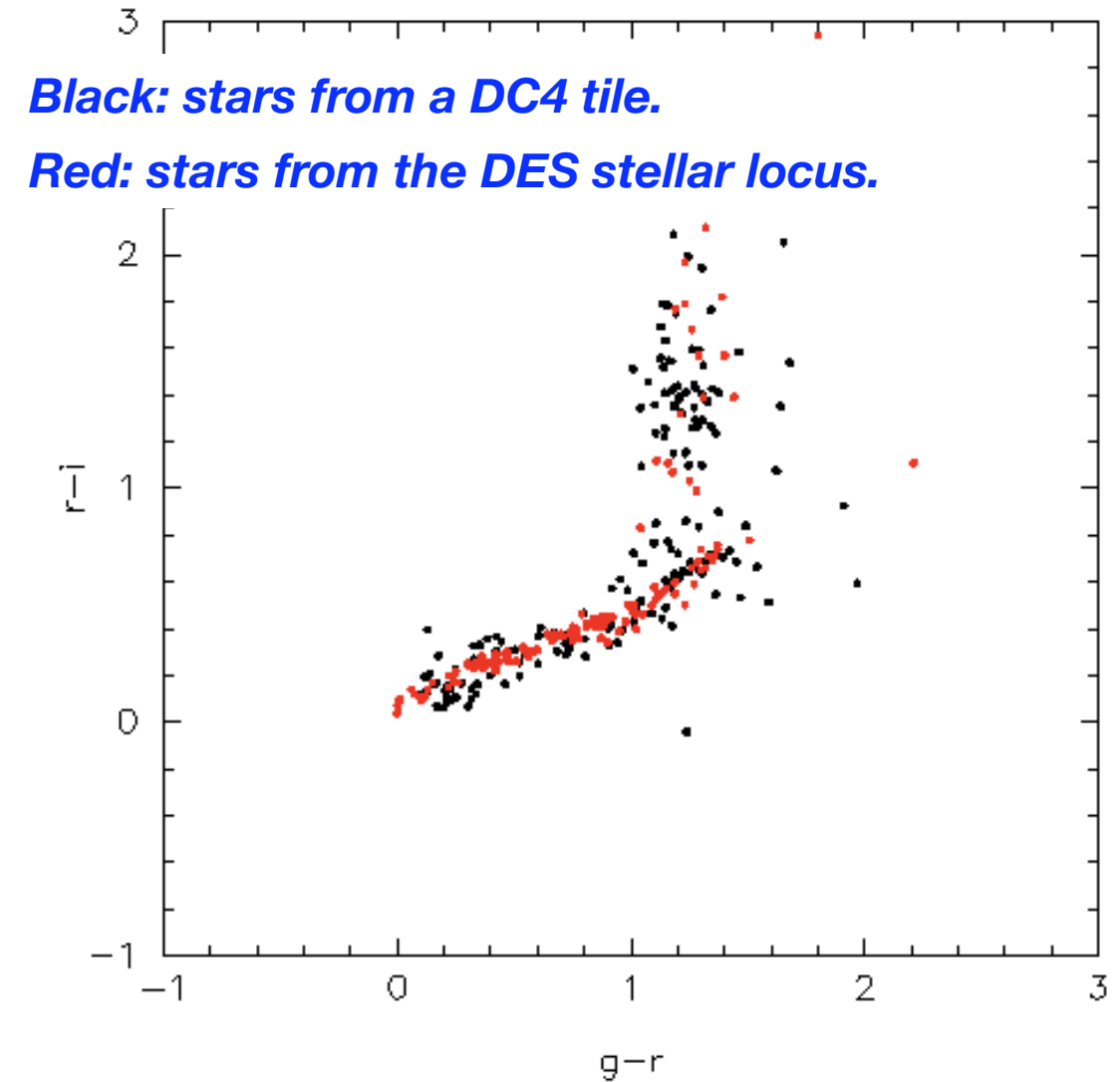
Photometric Calibration III

1. **2% photometric calibration required** ✗

   – **1% goal**

2. Fit DES/Gunn-Stryker locus to g-r, r-i, and i-z colors of 20<i<22 stars

   - allow (g-r), (r-i), (i-z) to vary

   - iterative gridded search for least-square residuals

   - final search on a 0.1% grid

3. Dispersion in offsets

- mag_auto	mag_aper1
g-r: 0.079	g-r: 0.160
r-i: 0.037	r-i: 0.117
i-z: 0.036	i-z: 0.109

4. These are calibration errors.

2234−3145

Black: stars from a DC4 tile.

Red: stars from the DES stellar locus.

Offsets for mag_auto g-r

$\sigma = 0.08$

Photometric Calibration IV

Stars in tile 2233-3438 at 20<i<22.

Red: mag_aper1

Black: mag_aper4

1. Aperture corrections are necessary for calculating total magnitudes.✕

2. Most of the calibration errors in mag_aper1 can be traced to variable seeing and the lack of aperture corrections.

3. This can induce large shifts in color.

   - 0.5 magnitude shift in g-r between the 1.5" aper1 and the 6.0" aper4.

4. For science, aperture corrections should be applied.

   - If it is good for star colors, it is good for galaxy colors.

   - This is true for all magnitudes.


2233–3438

r-i

g-r
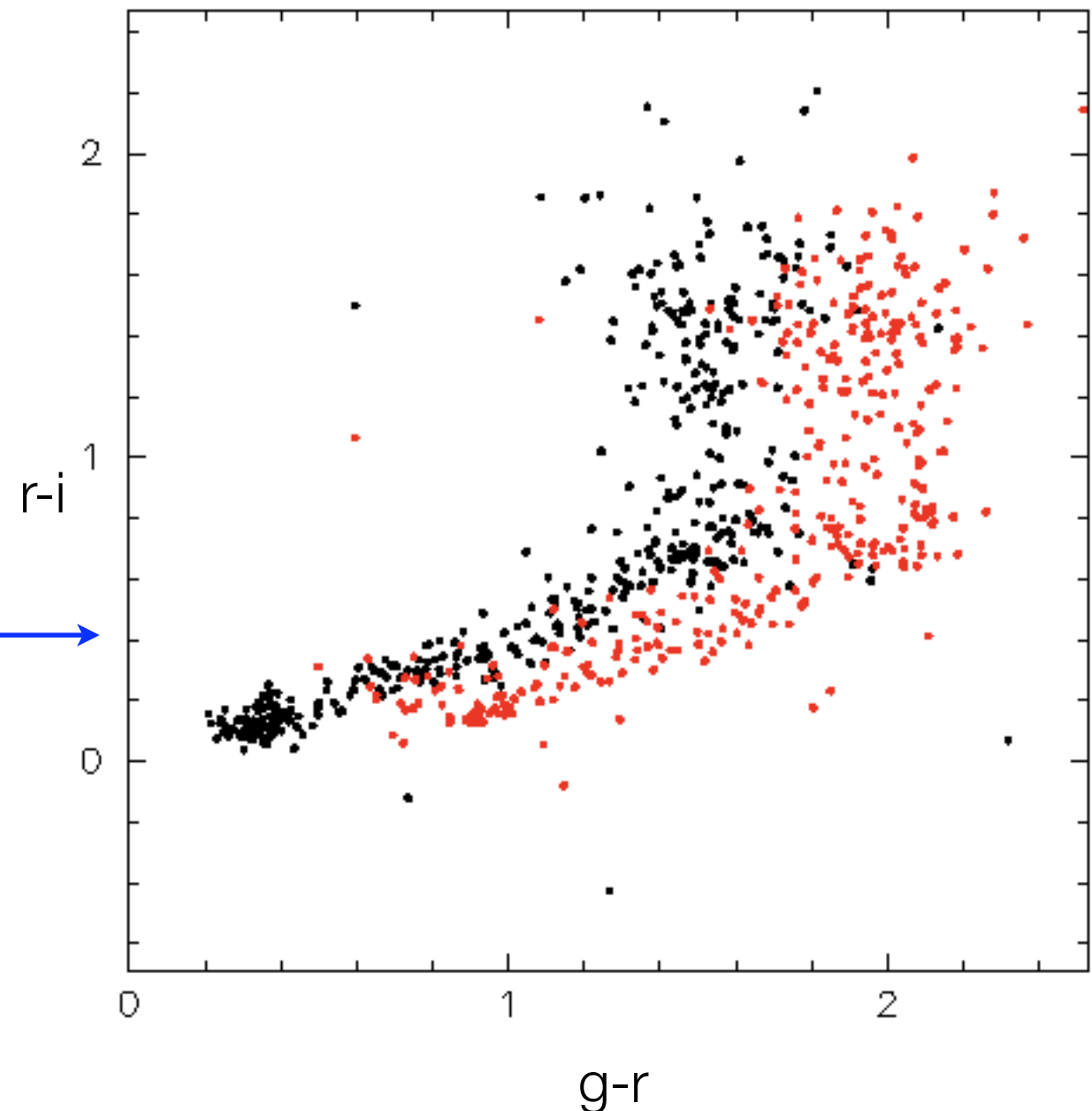
Magnitudes

1. Mag_auto performance

2. There is a magnitude dependent offset from the true magnitude.

   - amplitude is 0.05 mag/mag

   - counts against calibration budget

3. The residuals distribution is larger than I would expect.

   - At the 10σ magnitude i=22.8, the dispersion of mag-true is, after sigma-clipping, σ=0.27 mags.

   - There is a tail to negative residuals- as if mag_auto is over-reporting magnitudes.

     1. Matching errors in the testing?

     2. Deblending?

   - ~40% of objects unmatched to truth tables at i=22. Same as above?

2226−2601

mag_auto-true_mag

i-band mag_auto

22.7 < i < 22.9

% of objects

mag_auto-true_mag
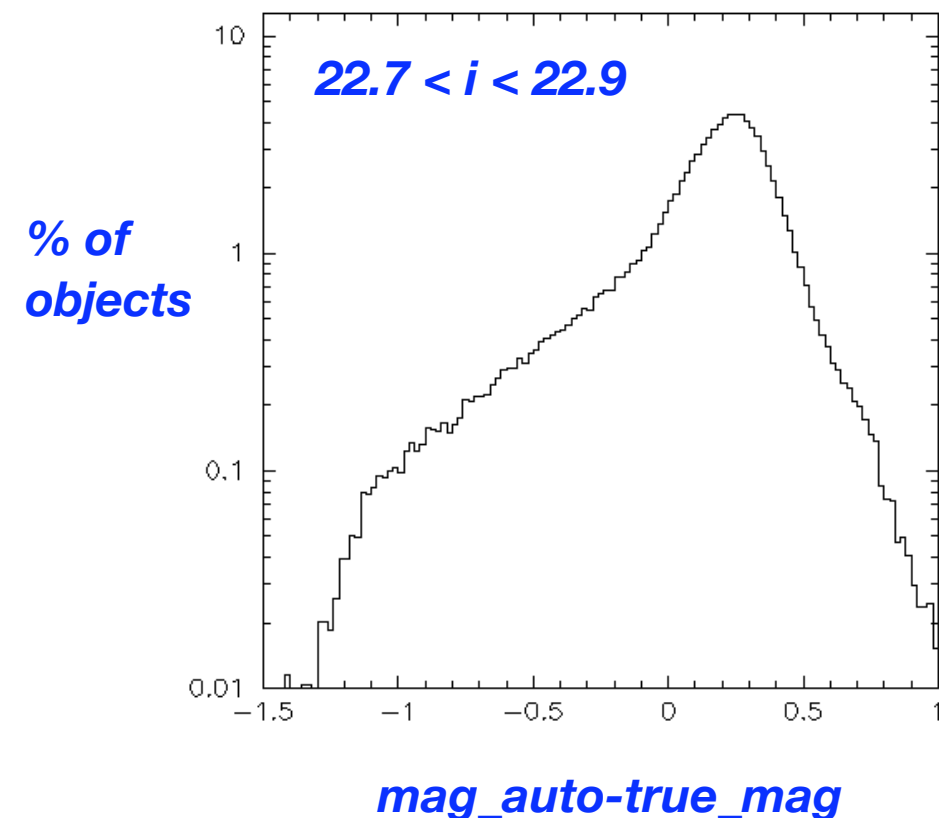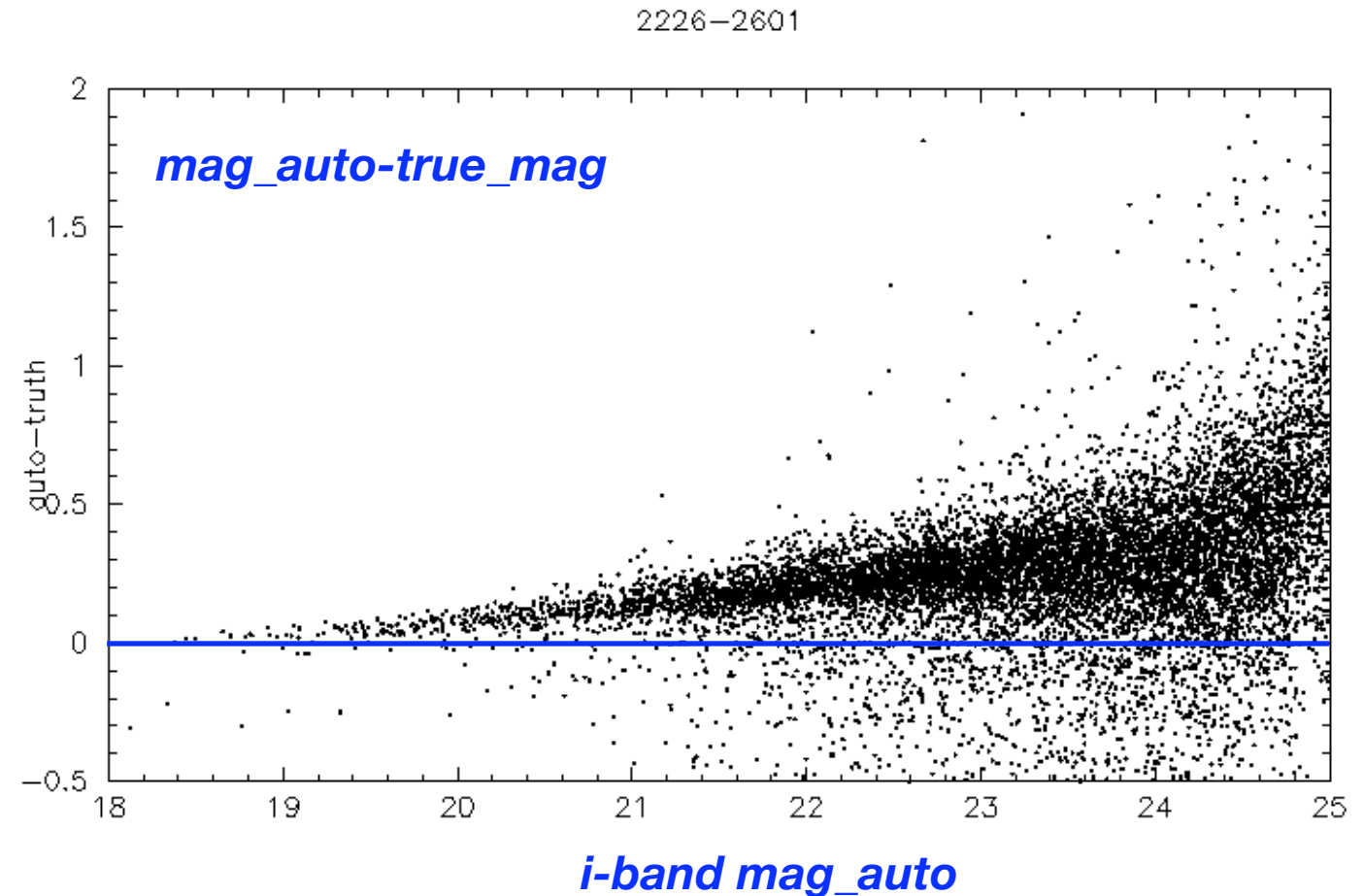
Photometric Redshifts

1. The photo-z dispersion averaged over all galaxies in the sample should be less than $\sigma_z = 0.12$. ✕

2. The photo-z are a catalog of neural net photo-zs. The solution was not optimized. The photo-z are compared against the true z from the mock catalog.

3. $\sigma_z > 0.12$ at most z.

4. In each bin of 0.1 in redshift from z=0 to z=1.5 the fraction of galaxies with photo-z errors larger than 2 (3) times the photo-z dispersion in that bin should be less than 10% (1.5%), and the first fraction above needs to be known within 1%. ✕

5. The tails are fairly well constrained. Here the 3σ tail is 1.7% of total population, slightly higher than the required 1.5%.
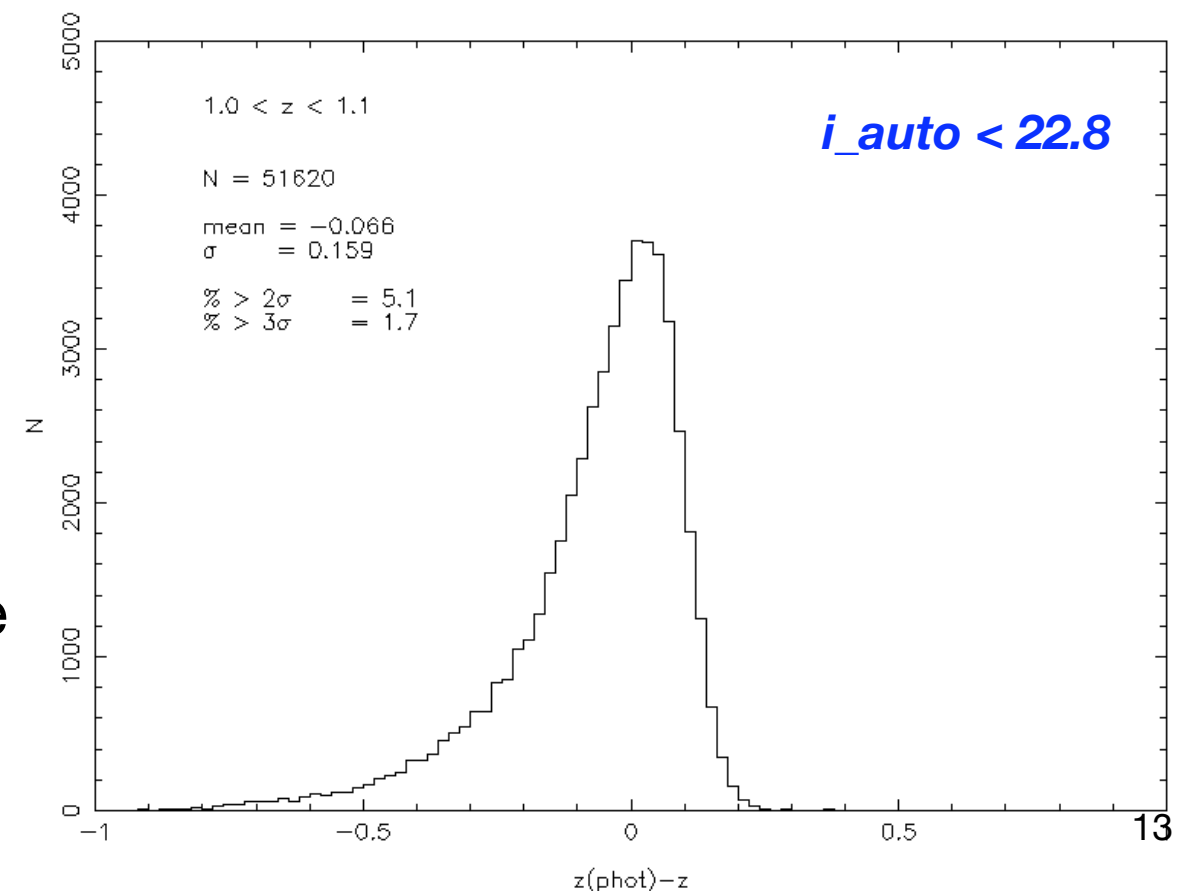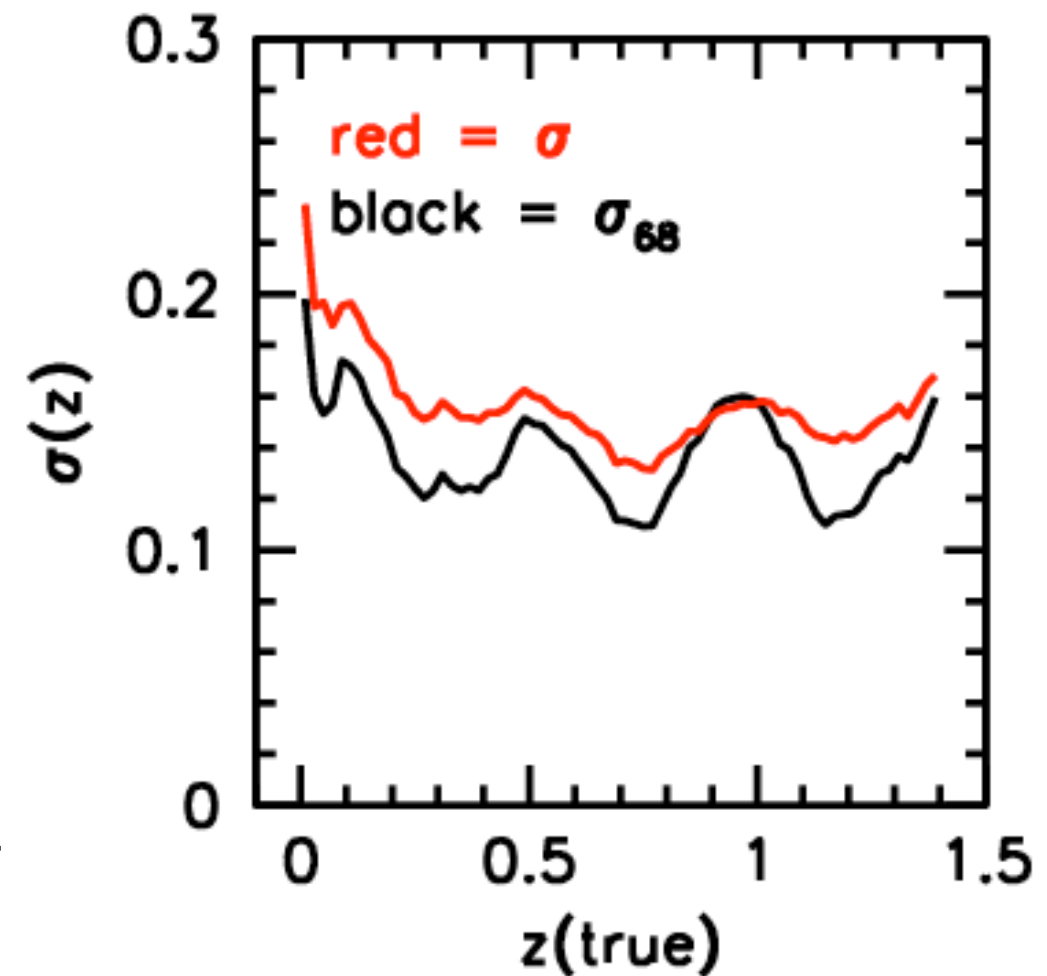
Summary of Tests

1. We report on a subset of the requirements
   - Image quality   ✗
   - Limiting magnitude   ✓✗
   - Astrometric calibration   ✓
   - Photometric calibration   ✗
      - Galaxy catalog content   ✗
   - Photometric redshifts   ✗

2. The testing is an iterative process.
   - Neither the simulation, the data reduction software, nor the tests are of the quality we expect at the time of scientific commissioning.

3. Furthermore:
   - A signed off DES Science Requirements Document ✗

Science Commissioning Tests

1. Reported here are tests of DC4 against the SRD.
   - Currently, the process points out places where the simulation and the data reduction codes can be improved.

2. Many of the commissioning tests of the DES system will be identical to SRD testing of the data challenges
   - These tests are the natural core QA tests- more can be abstracted from SWG work
   - These same tests are the natural core of the science commissioning tests.

3. During commissioning it will be of value to have a deep understanding of the data reduction system. The data reduction system then becomes a tool to understand the performance of the real camera and the actual telescope.
   - Commissioning involves the DECam system, the DESDM system, and the E2E system/Community Pipeline.
   - The on-sky phase of commissioning concludes after acceptance tests that include relevant performance requirements derived from the SRD and Community Needs documents.

4. Impetus to think about commissioning: plans that must be developed over the next year.
   - DECam Installation and Commissioning Plan
   - DESDM Commissioning Plan
   - DECam/E2E System Integration and Commissioning Plan
   - Preliminary Observing Plan

5. In the long run, these tests are needed to assure that the data quality will continue to meet the DES science goals after scientific commissioning is completed.

Data quality is paramount. We must test.

Backup Slides

Data Access

1. We access the data at the secondary archive

   - files and directories on disk

   - The SWGs are testing the portal interface.

2. The catalogs of the tiles - 0.75x0.75 degree

   - Directories labeled by processing date

     1. Guidance from DESDM is to use the latest processing date.

     2. We adopt the 2008-12-* processing

        - Eliminate multiple processings near this date

        - Eliminate dates without coadd dirs

   - Eliminate everything at Dec > -20

   - 233 tiles in final list.
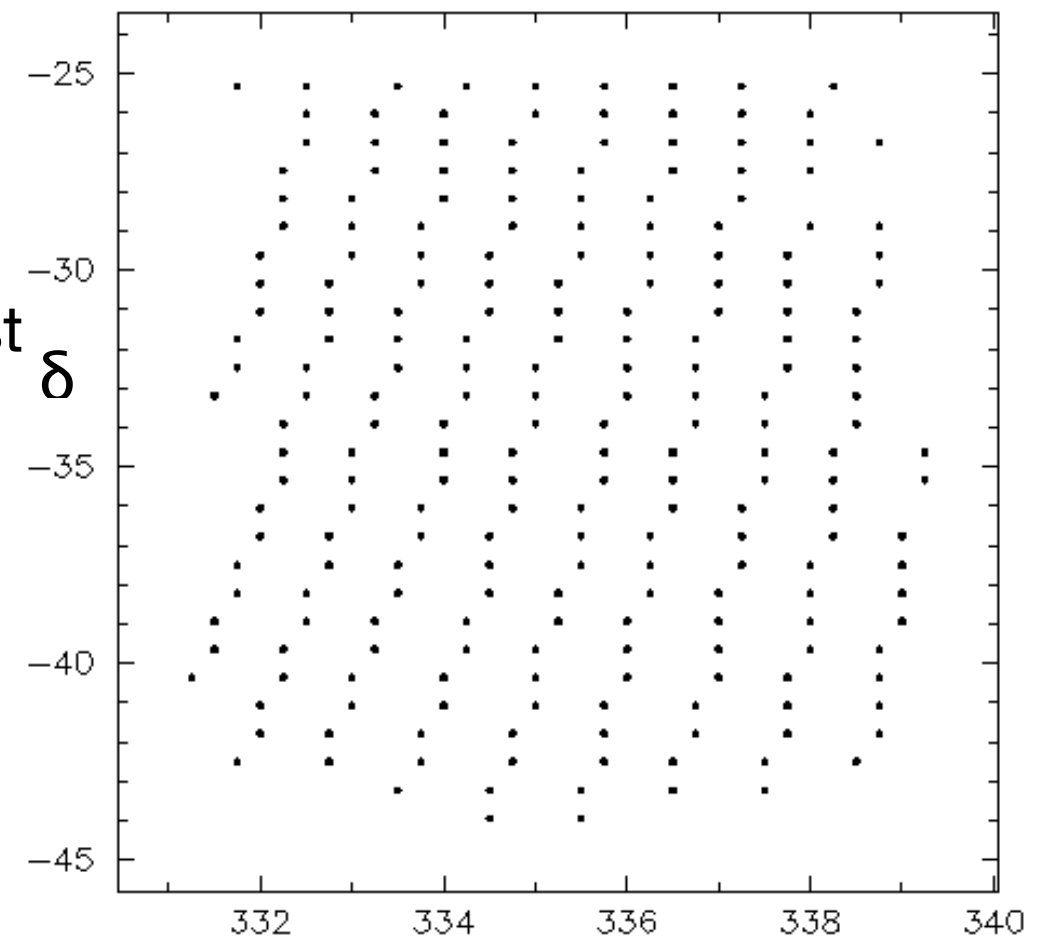
The location of the 233 tiles on the sky.

Overlapping tiles

Black- tile 1   Green- tile 2   Orange- tile 3

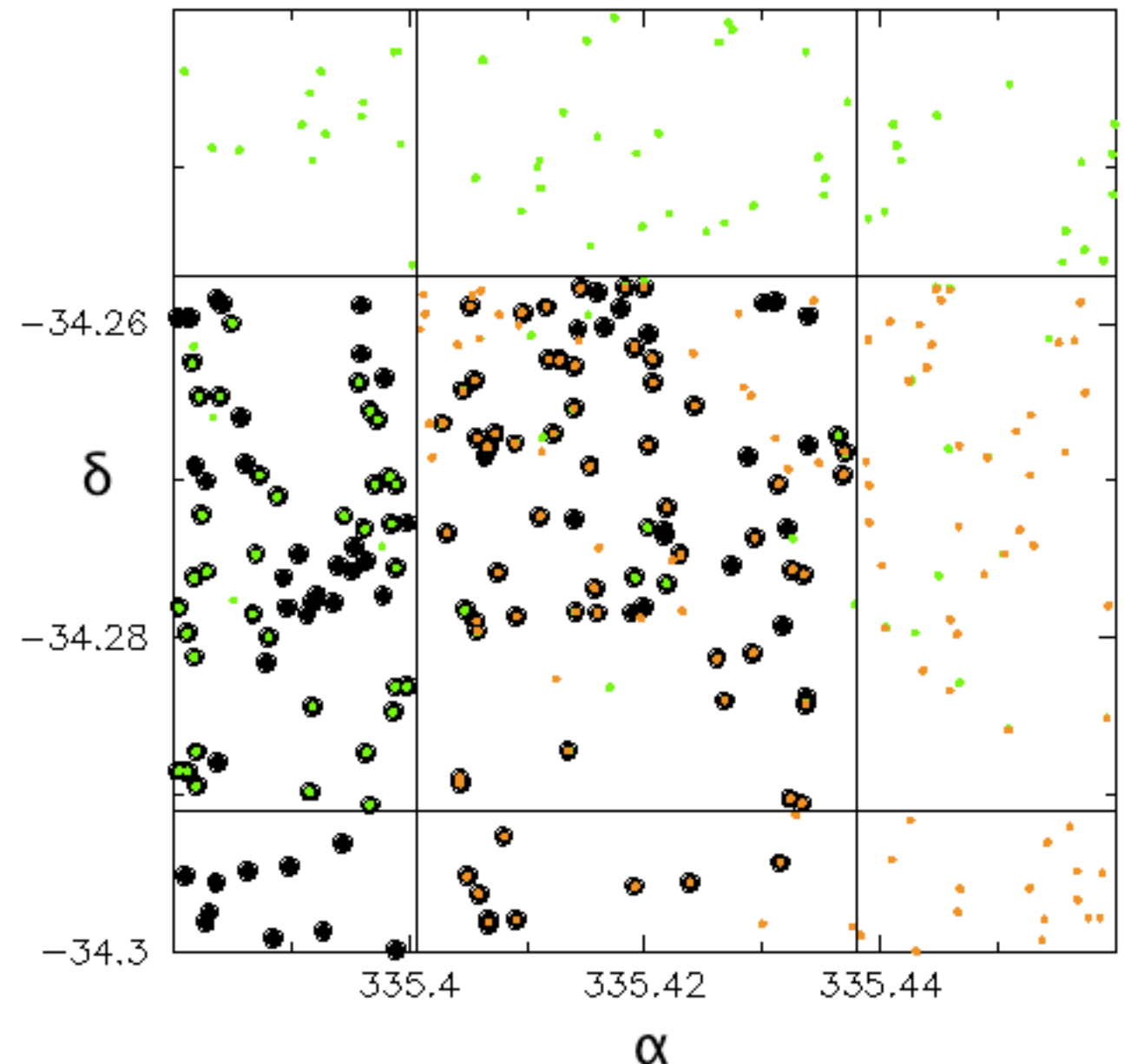1. The tiles have overlapping boundaries.

- What is the algorithm to reject duplicate objects?

  1. If geometric, the cuts is not uniform in RA and Dec.

  2. If RA, Dec matching then objects from different tiles contribute to the final catalog

- We assume a 1' overlap on every side.

A corner of overlapping tiles. The lines show the approximate boundaries of the objects in the individual tiles. Note the rectangular shape- this is not made square by a cos(δ) correction. What size is the geometrical cutout?

Photometric Calibration

The universe in the DC4 simulation is different from that in the real universe.

Stars in DC4 (right) at 20<i<22 have properties similar to the real universe (left).

20 < i < 22 is used for testing purposes.

Stars at i<19 are from USNO-B and the scatter in the colors reflects the photometric scatter in the USNO-B catalog.

Stars at i>19 are from the Besacon model and use Besacon SDSS colors.

m stars

thick disk

halo

g-r

SDSS Stripe 82

i

2218-3646

g-r

2218-3646

i
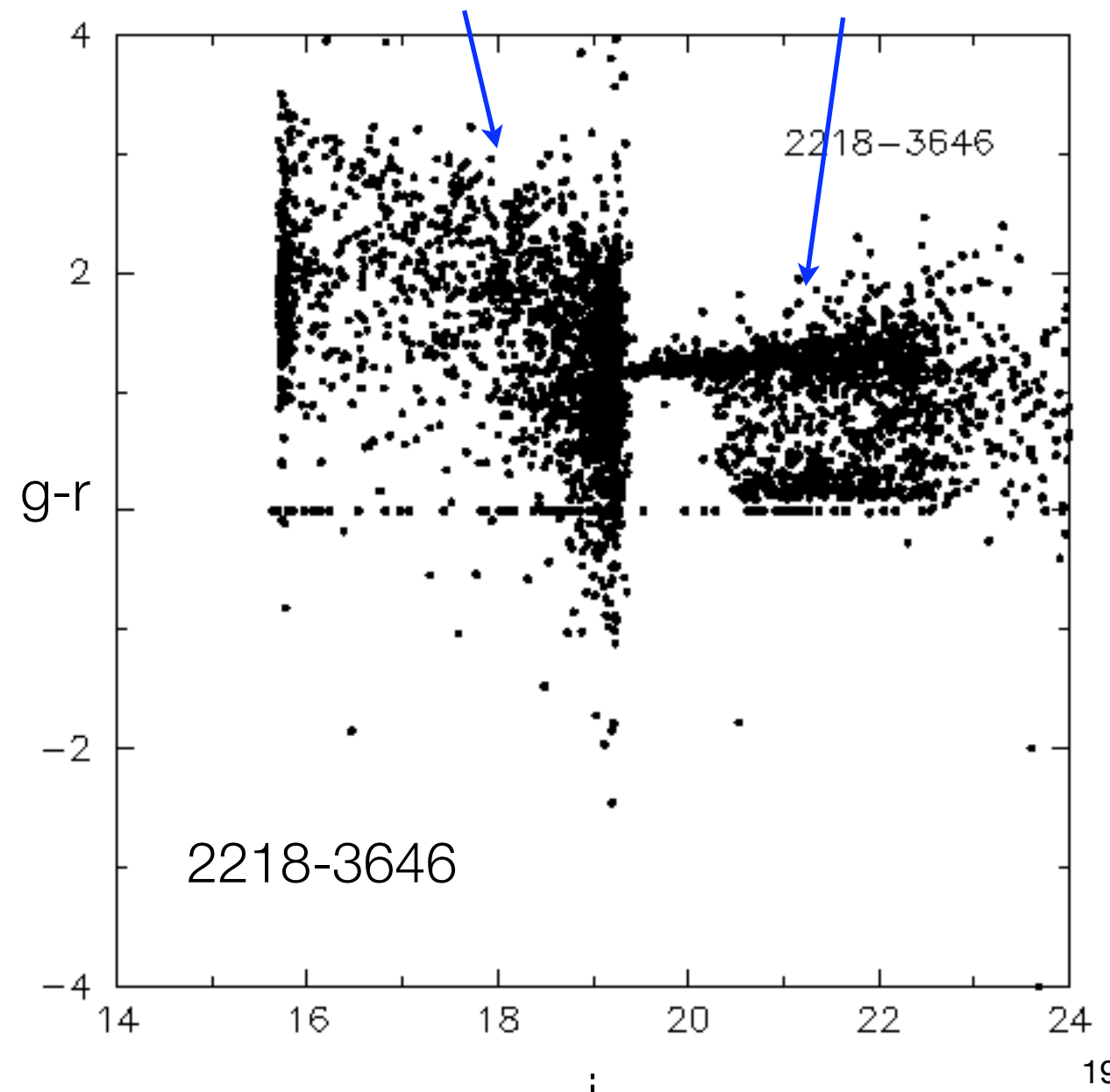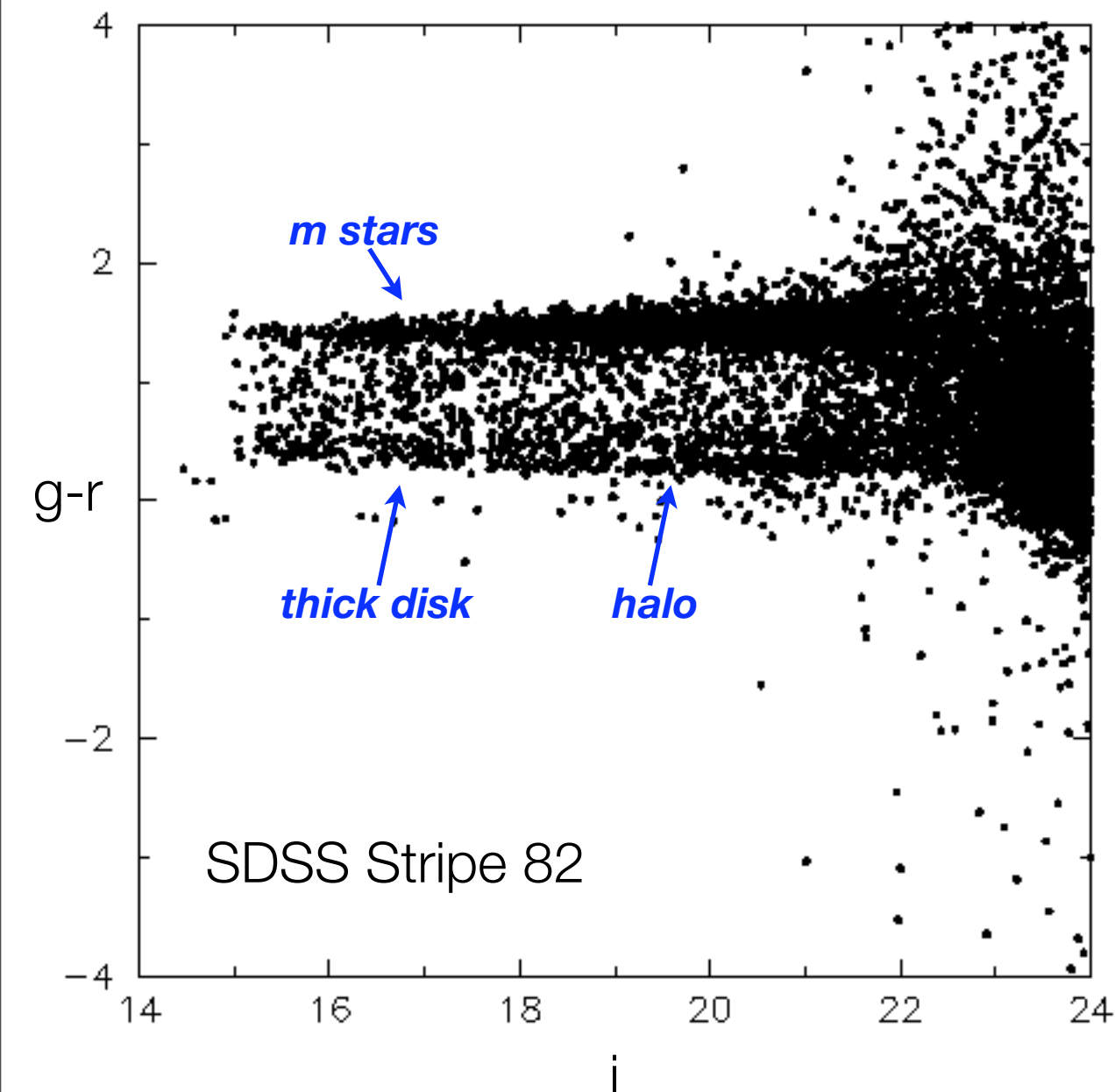
Stars in DC5

1. We have asked the Galactic Structure Study Group to perform the star simulation for DC5

   - USNO-B stars at r < 20

     1. Use USNO-B colors to choose an SED

     2. Convolve SED with DES system response

   - Santiago+Yanny galactic structure model for r > 20

     1. Stellar SEDs convolved with DES system response

2. The Science Committee and the Co-coordinators of the study group have agreed.

3. This allows the Galactic Structure study group to emplace interesting science into the simulation for their group to try to extract.

Shifted Objects?

Is the 40% impurity deblending or matching issues? Unknown.

(333.4639, -38.2822)
OMAG(i) = 22.5

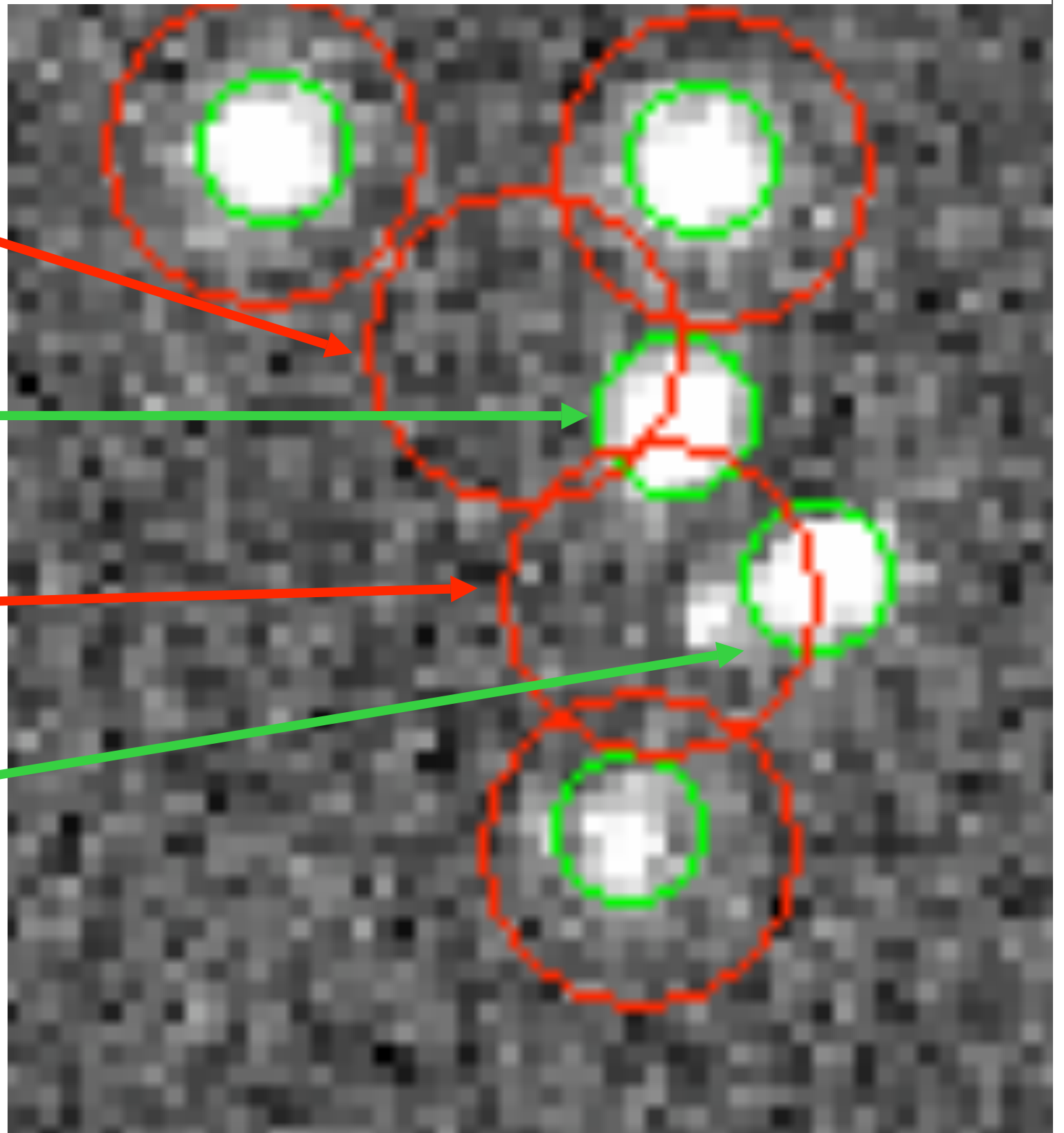(333.4642, -38.2821)
MAG_AUTO = 22.1

(333.4655, -38.2813)
OMAG(i) = 22.3

(333.4648, -38.2616)
MAG_AUTO = 21.2
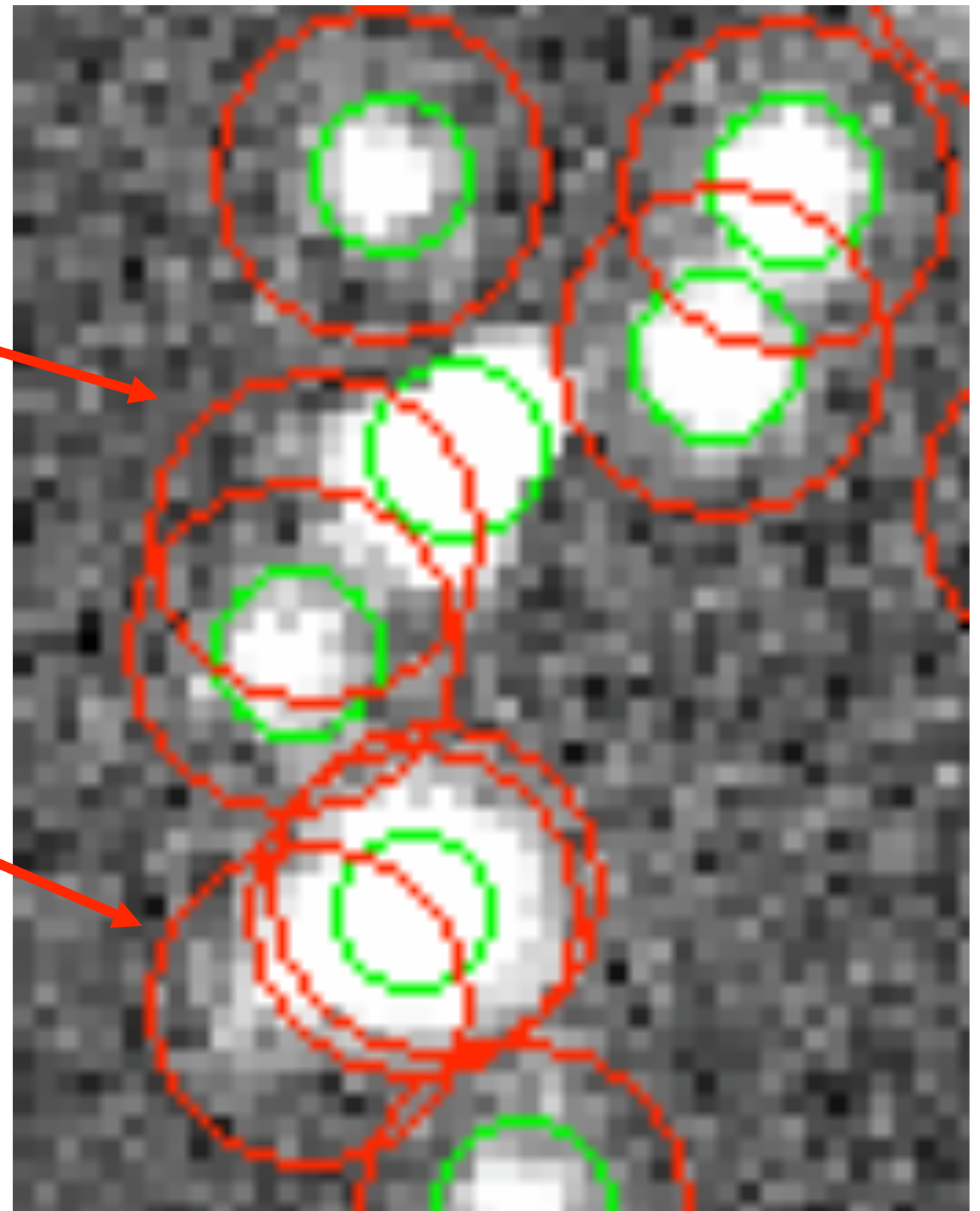
Red=CatSim6 (input)
Green = DC4 (output)

Vic Scarpine

Deblending/ Missed Objects?

Is the 40% impurity deblending or matching issues? Unknown.

(333.4620, -38.2822)
OMAG(i) = 20.4

(333.4620, -38.2838)
OMAG(i) = 20.8

Red=CatSim6 (input)
Green = DC4 (output)

Vic Scarpine

Deblending Mess?

MAG_AUTO(i) =19.0

MAG_AUTO(i) =19.2

Red=CatSim6
Green = DC4
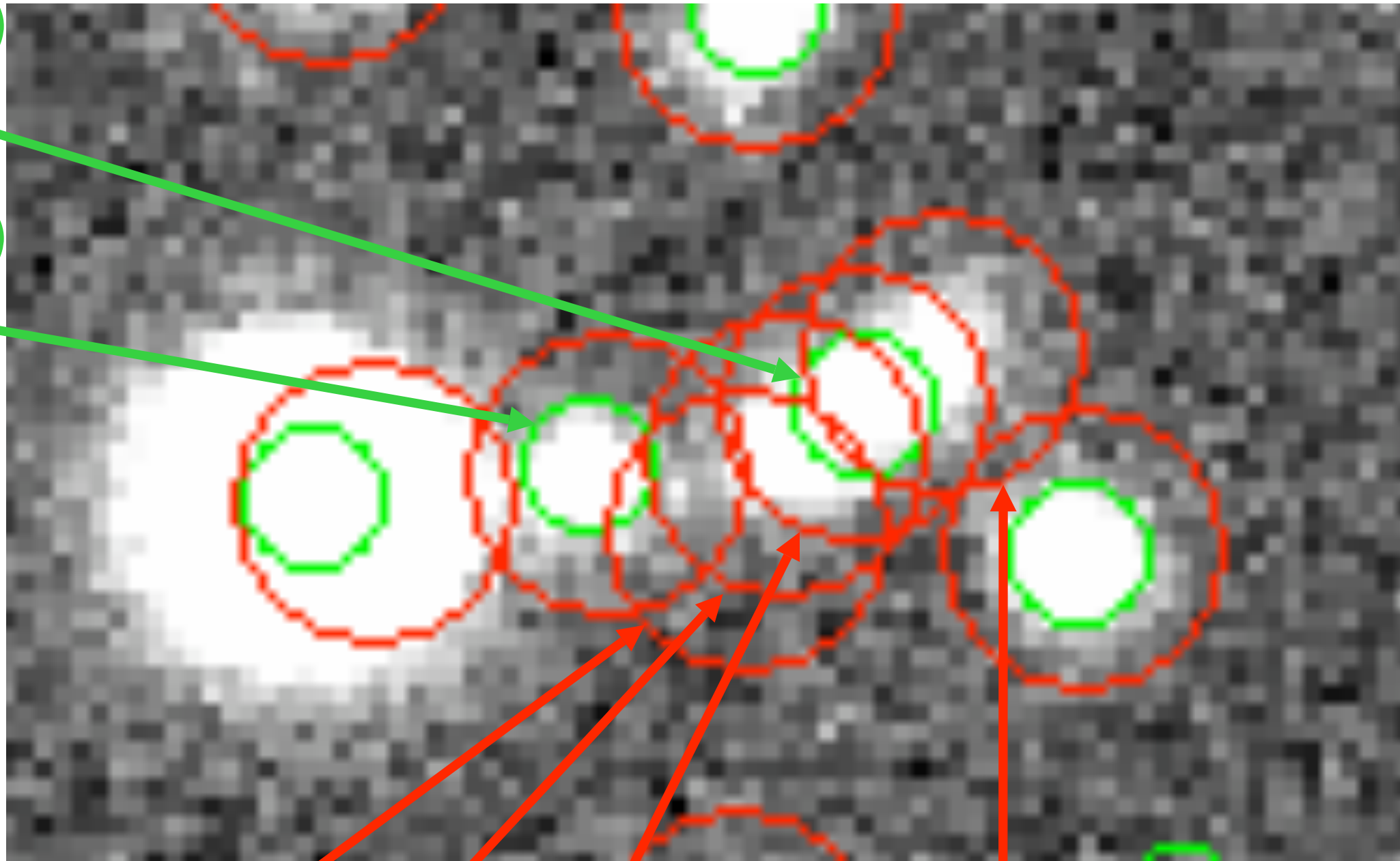
(333.4589,-38.2882)
OMAG(i) = 22.7

(333.4587,-38.2879)
OMAG(i) = 22.1

(333.4583,-38.2877)
OMAG(i) = 22.4

(333.4579,-38.2875)
OMAG(i) = 21.8